



13 OIPE
Date
6/26/02

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/081,218

DATE: 03/12/2002
TIME: 09:57:00

Input Set : N:\Crf3\RULE60\10081218.raw
Output Set: N:\CRF3\03122002\J081218.raw

SEQUENCE LISTING
ENTERED

3 (1) GENERAL INFORMATION:
5 (i) APPLICANT: SHIBAYAMA, SHIRO
6 HIRANO, ATSUSHI
7 OHNO, HIROYUKI
9 (ii) TITLE OF INVENTION: A NOVEL POLYPEPTIDE AND DNAS ENCODING IT
11 (iii) NUMBER OF SEQUENCES: 6
13 (iv) CORRESPONDENCE ADDRESS:
14 (A) ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
15 (B) STREET: 2100 Pennsylvania Avenue, N.W.
16 (C) CITY: Washington
17 (D) STATE: D.C.
18 (E) COUNTRY: U.S.A.
19 (F) ZIP: 20037-3202
21 (v) COMPUTER READABLE FORM:
22 (A) MEDIUM TYPE: Floppy disk
23 (B) COMPUTER: IBM PC compatible
24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
27 (vi) CURRENT APPLICATION DATA:
C--> 28 (A) APPLICATION NUMBER: US/10/081,218
C--> 29 (B) FILING DATE: 25-Feb-2002
30 (C) CLASSIFICATION:
32 (vii) PRIOR APPLICATION DATA:
33 (A) APPLICATION NUMBER: US/09/246,355
34 (B) FILING DATE:
36 (A) APPLICATION NUMBER: US/08/852,811
37 (B) FILING DATE: 07-MAY-1997
39 (A) APPLICATION NUMBER: 08/439,457
40 (B) FILING DATE:
42 (A) APPLICATION NUMBER: JP 123155/1994
43 (B) FILING DATE: 12-MAY-1994
45 (ix) TELECOMMUNICATION INFORMATION:
46 (A) TELEPHONE: (202)293-7060
47 (B) TELEFAX: (202)293-7860
48 (C) TELEX: 6491103
51 (2) INFORMATION FOR SEQ ID NO: 1:
53 (i) SEQUENCE CHARACTERISTICS:
54 (A) LENGTH: 205 amino acids
55 (B) TYPE: amino acid
56 (D) TOPOLOGY: linear
58 (ii) MOLECULE TYPE: protein
63 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

RECEIVED
APR 18 2002
TECH CENTER 1600/2900

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/081,218

DATE: 03/12/2002
TIME: 09:57:00

Input Set : N:\Crf3\RULE60\10081218.raw
Output Set: N:\CRF3\03122002\J081218.raw

66 Met Gly Thr Ala Gly Ala Met Gln Leu Cys Trp Val Ile Leu Gly Phe
67 -24 -20 -15 -10
68 Leu Leu Phe Arg Gly His Asn Ser Gln Pro Thr Met Thr Gln Thr Ser
69 -5 1 5
71 Ser Ser Gln Gly Gly Leu Gly Leu Ser Leu Thr Thr Glu Pro Val
72 10 15 20
74 Ser Ser Asn Pro Gly Tyr Ile Pro Ser Ser Glu Ala Asn Arg Pro Ser
75 25 30 35 40
77 His Leu Ser Ser Thr Gly Thr Pro Gly Ala Gly Val Pro Ser Ser Gly
78 45 50 55
80 Arg Asp Gly Gly Thr Ser Arg Asp Thr Phe Gln Thr Val Pro Pro Asn
81 60 65 70
83 Ser Thr Thr Met Ser Leu Ser Met Arg Glu Asp Ala Thr Ile Leu Pro
84 75 80 85
86 Ser Pro Thr Ser Glu Thr Val Leu Thr Val Ala Ala Phe Gly Val Ile
87 90 95 100
89 Ser Phe Ile Val Ile Leu Val Val Val Val Ile Ile Leu Val Gly Val
90 105 110 115 120
92 Val Ser Leu Arg Phe Lys Cys Arg Lys Ser Lys Glu Ser Glu Asp Pro
93 125 130 135
95 Gln Lys Pro Gly Ser Ser Gly Leu Ser Glu Ser Cys Ser Thr Ala Asn
96 140 145 150
98 Gly Glu Lys Asp Ser Ile Thr Leu Ile Ser Met Lys Asn Ile Asn Met
99 155 160 165
101 Asn Asn Gly Lys Gln Ser Leu Ser Ala Glu Lys Val Leu
102 170 175 180
104 (2) INFORMATION FOR SEQ ID NO: 2:
106 (i) SEQUENCE CHARACTERISTICS:
107 (A) LENGTH: 618 base pairs
108 (B) TYPE: nucleic acid
109 (C) STRANDEDNESS: single
110 (D) TOPOLOGY: linear
112 (ii) MOLECULE TYPE: cDNA to mRNA
117 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
119 ATGGGCACCG CAGGAGCCAT GCAGCTGTGC TGGGTGATCC TGGGCTTCCT CCTGTTCCGA 60
121 GGCCACAACCT CCCAGCCAC AATGACCCAG ACCTCTAGCT CTCAGGGAGG CCTTGGCGGT 120
123 CTAAGTCTGA CCACAGAGCC AGTTTCTTCC AACCCAGGAT ACATCCCTTC CTCAGAGGCT 180
125 AACAGGCCAA GCCATCTGTC CAGCACTGGT ACCCCAGGCG CAGGTGTCCC CAGCAGTGG 240
127 AGAGACGGAG GCACAAGCAG AGACACATT CAAACTGTTC CCCCCAATTC ACCACCATG 300
129 AGCCTGAGCA TGAGGGAAGA TGCGACCATC CTGCCAGCC CCACGTCAGA GACTGTGCTC 360
131 ACTGTGGCTG CATTGGTGT TATCAGCTC ATTGTCATCC TGGTGGTTGT GGTGATCATC 420
133 CTAGTTGGTG TGGTCAGCCT GAGGTTCAAG TGTCCGAAG GCAAGGAGTC TGAAGATCCC 480
135 CAGAACCTG GGAGTCAGG GCTGTCTGAA AGCTGCTCCA CAGCCAATGG AGAGAAAGAC 540
137 AGCATCACCC TTATCTCCAT GAAGAACATC AACATGAATA ATGGCAAACA AAGTCTCTCA 600
139 GCAGAGAAGG TTCTTTAA 618

141 (2) INFORMATION FOR SEQ ID NO: 3:
143 (i) SEQUENCE CHARACTERISTICS:
144 (A) LENGTH: 983 base pairs
145 (B) TYPE: nucleic acid

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/081,218

DATE: 03/12/2002
 TIME: 09:57:00

Input Set : N:\Crf3\RULE60\10081218.raw
 Output Set: N:\CRF3\03122002\J081218.raw

146 (C) STRANDEDNESS: single
 147 (D) TOPOLOGY: linear
 149 (ii) MOLECULE TYPE: cDNA to mRNA
 154 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 156 GCCTGCCCG CACATACCCA GCTGACATGG GCACCGCAGG AGCCATGCAG CTGTGCTGGG 60
 158 TGATCCTGGG CTTCTCCTG TTCCGAGGCC ACAACTCCA GCCCACAAATG ACCCAGACCT 120
 160 CTAGCTCTCA GGGAGGCCTT GGCGGTCTAA GTCTGACCAC AGAGCCAGTT TCTTCCAACC 180
 162 CAGGATACAT CCCTCCTCA GAGGCTAACAA GGCCAAGCCA TCTGTCCAGC ACTGGTACCC 240
 164 CAGGCGCAGG TGTCCCCAGC AGTGGAAAGAG ACGGAGGCAC AAGCAGAGAC ACATTTCAA 300
 166 CTGTTCCCCC CAATTCACCC ACCATGAGCC TGAGCATGAG GGAAGATGCG ACCATCCTGC 360
 168 CCAGCCCCAC GTCAAGAGACT GTGCTCACTG TGGCTGCATT TGGTGTATC AGCTTCATTG 420
 170 TCATCCTGGT GGTGTGGTG ATCATCCTAG TTGGTGTGGT CAGCCTGAGG TTCAAGTGTC 480
 172 GGAAGAGCAA GGAGTCTGAA GATCCCCAGA AACCTGGGAG TTCAGGGCTG TCTGAAAGCT 540
 174 GCTCCACAGC CAATGGAGAG AAAGACAGCA TCACCCCTTAT CTCCATGAAG AACATCAACA 600
 176 TGAATAATGG CAAACAAAGT CTCTCAGCAG AGAAGGTTCT TTAAAAGCAA CTTTGGGTCC 660
 178 CCATGAGTCC AAGGATGATG CAGCTGCCCT GTGACTACAA GGAGGAAGAG ATGGAATTAG 720
 180 TAGAGGCAAT GAACCACATG TAAATTATTT TATTGTTCA TGTCTGCTTC TAGATCTAAA 780
 182 GGACACTAGC ATTGCCCGAG ATCTGGGAGC AAGCTACCAA CAGGGGAGAC TCTTCCTGT 840
 184 ATGGACAGCT GCTGTGGAAA TACTGCCTGC TTCTCCCACC TCCTCAGAGC CACAGGAAAG 900
 186 AGGAGGTGAC AGAGAGAGAG CAAGGAAAGT GATGAGGTGG ATTGATACTT TCTACTTTGC 960
 188 ATTAAAATTA TTTCTAGCC TGC 983
 190 (2) INFORMATION FOR SEQ ID NO: 4:
 192 (i) SEQUENCE CHARACTERISTICS:
 193 (A) LENGTH: 983 base pairs
 194 (B) TYPE: nucleic acid
 195 (C) STRANDEDNESS: single
 196 (D) TOPOLOGY: linear
 198 (ii) MOLECULE TYPE: cDNA to mRNA
 200 (vi) ORIGINAL SOURCE:
 201 (A) ORGANISM: Homo sapiens
 202 (H) CELL LINE: endothelial cell line of umbilical cord vein
 204 (ix) FEATURE:
 205 (A) NAME/KEY: CDS
 206 (B) LOCATION: 27..644
 207 (C) IDENTIFICATION METHOD: by similarity to some other pattern
 209 (ix) FEATURE:
 210 (A) NAME/KEY: sig_peptide
 211 (B) LOCATION: 27..98
 212 (C) IDENTIFICATION METHOD: by similarity with known sequence or
 213 to an established consensus
 215 (ix) FEATURE:
 216 (A) NAME/KEY: misc_feature
 217 (B) LOCATION: 381..464
 218 (C) IDENTIFICATION METHOD: by similarity with known sequence or
 219 to an established consensus
 221 (ix) FEATURE:
 222 (A) NAME/KEY: mat_peptide
 223 (B) LOCATION: 99..641
 224 (C) IDENTIFICATION METHOD: by similarity with known sequence or

RAW SEQUENCE LISTING DATE: 03/12/2002
 PATENT APPLICATION: US/10/081,218 TIME: 09:57:00

Input Set : N:\Crf3\RULE60\10081218.raw
 Output Set: N:\CRF3\03122002\J081218.raw

225 to an established consensus
 227 (ix) FEATURE:
 228 (A) NAME/KEY: misc_feature
 229 (B) LOCATION: 312..320
 230 (C) IDENTIFICATION METHOD: by similarity with known sequence or
 231 to an established consensus
 234 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 236 GCCTGCCCGC CACATACCCA GCTGAC ATG GGC ACC GCA GGA GCC ATG CAG CTG 53
 237 Met Gly Thr Ala Gly Ala Met Gln Leu
 238 -24 -20
 240 TGC TGG GTG ATC CTG GGC TTC CTC CTG TTC CGA GGC CAC AAC TCC CAG 101
 241 Cys Trp Val Ile Leu Gly Phe Leu Leu Phe Arg Gly His Asn Ser Gln
 242 -15 -10 -5 1
 244 CCC ACA ATG ACC CAG ACC TCT AGC TCT CAG GGA GGC CTT GGC GGT CTA 149
 245 Pro Thr Met Thr Gln Thr Ser Ser Gln Gly Gly Leu Gly Gly Leu
 246 5 10 15
 248 AGT CTG ACC ACA GAG CCA GTT TCT TCC AAC CCA GGA TAC ATC CCT TCC 197
 249 Ser Leu Thr Thr Glu Pro Val Ser Ser Asn Pro Gly Tyr Ile Pro Ser
 250 20 25 30
 252 TCA GAG GCT AAC AGG CCA AGC CAT CTG TCC AGC ACT GGT ACC CCA GGC 245
 253 Ser Glu Ala Asn Arg Pro Ser His Leu Ser Ser Thr Gly Thr Pro Gly
 254 35 40 45
 256 GCA GGT GTC CCC AGC AGT GGA AGA GAC GGA GGC ACA AGC AGA GAC ACA 293
 257 Ala Gly Val Pro Ser Ser Gly Arg Asp Gly Gly Thr Ser Arg Asp Thr
 258 50 55 60 65
 260 TTT CAA ACT GTT CCC CCC AAT TCA ACC ACC ATG AGC CTG AGC ATG AGG 341
 261 Phe Gln Thr Val Pro Pro Asn Ser Thr Thr Met Ser Leu Ser Met Arg
 262 70 75 80
 264 GAA GAT GCG ACC ATC CTG CCC AGC CCC ACG TCA GAG ACT GTG CTC ACT 389
 265 Glu Asp Ala Thr Ile Leu Pro Ser Pro Thr Ser Glu Thr Val Leu Thr
 266 85 90 95
 268 GTG GCT GCA TTT GGT GTT ATC AGC TTC ATT GTC ATC CTG GTG GTT GTG 437
 269 Val Ala Ala Phe Gly Val Ile Ser Phe Ile Val Ile Leu Val Val Val
 270 100 105 110
 272 GTG ATC ATC CTA GTT GGT GTG GTC AGC CTG AGG TTC AAG TGT CGG AAG 485
 273 Val Ile Ile Leu Val Gly Val Val Ser Leu Arg Phe Lys Cys Arg Lys
 274 115 120 125
 276 AGC AAG GAG TCT GAA GAT CCC CAG AAA CCT GGG AGT TCA GGG CTG TCT 533
 277 Ser Lys Glu Ser Glu Asp Pro Gln Lys Pro Gly Ser Ser Gly Leu Ser
 278 130 135 140 145
 280 GAA AGC TGC TCC ACA GCC AAT GGA GAG AAA GAC AGC ATC ACC CTT ATC 581
 281 Glu Ser Cys Ser Thr Ala Asn Gly Glu Lys Asp Ser Ile Thr Leu Ile
 282 150 155 160
 284 TCC ATG AAG AAC ATC AAC ATG AAT AAT GGC AAA CAA AGT CTC TCA GCA 629
 285 Ser Met Lys Asn Ile Asn Met Asn Asn Gly Lys Gln Ser Leu Ser Ala
 286 165 170 175
 288 GAG AAG GTT CTT TAA AAGCAACTTT GGGTCCCCAT GAGTCCAAGG ATGATGCAGC 684
 289 Glu Lys Val Leu *
 290 180

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/081,218

DATE: 03/12/2002
 TIME: 09:57:00

Input Set : N:\Crf3\RULE60\10081218.raw
 Output Set: N:\CRF3\03122002\J081218.raw

292	TGCCCTGTGA CTACAAGGAG GAAGAGATGG AATTAGTAGA GGCAATGAAC CACATGTAAA	744
294	TTATTTATT GTTTCATGTC TGCTTCTAGA TCTAAAGGAC ACTAGCATTG CCCCAGATCT	804
296	GGGAGCAAGC TACCAACAGG GGAGACTCTT TCCTGTATGG ACAGCTGCTG TGGAAATACT	864
298	GCCTGCTTCT CCCACCTCCT CAGAGCCACA GGAAAGAGGA GGTGACAGAG AGAGAGCAAG	924
300	GAAAGTGATG AGGTGGATTG ATACTTTCTA CTTTGCATTA AAATTATTTT CTAGCCTGC	983

302 (2) INFORMATION FOR SEQ ID NO: 5:

304 (i) SEQUENCE CHARACTERISTICS:
 305 (A) LENGTH: 56 base pairs
 306 (B) TYPE: nucleic acid
 307 (C) STRANDEDNESS: double
 308 (D) TOPOLOGY: linear

W--> 310 (ii) MOLECULE TYPE: DNA (synthetic)

315 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

317	GGAGCTCGTC GACAGATCTG AATTCCATAT GCCCGGGGCG GCCGCCTAG TGGTAC	56
-----	--	----

319 (2) INFORMATION FOR SEQ ID NO: 6:

321 (i) SEQUENCE CHARACTERISTICS:
 322 (A) LENGTH: 56 base pairs
 323 (B) TYPE: nucleic acid
 324 (C) STRANDEDNESS: double
 325 (D) TOPOLOGY: linear

W--> 327 (ii) MOLECULE TYPE: DNA (synthetic)

332 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

334	CACTAGTGCG GCCGCCCCGG GCATATGGAA TTCAGATCTG TCGACGAGCT CCTGCA	56
-----	---	----

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/081,218

DATE: 03/12/2002

TIME: 09:57:01

Input Set : N:\Crf3\RULE60\10081218.raw
Output Set: N:\CRF3\03122002\J081218.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:310 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=5
L:327 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=6